

Strengths and Difficulties Questionnaire Parent Report (SDQ-PR) as Screening Instrument of Children Mental Health in Indonesia

Supra Wimbar¹, Juliarni Siregar², Mistety Oktaviana³, & Regi Astriningsih⁴

Fakultas Psikologi Universitas Gadjah Mada

Abstract. This study aimed at examining the agreement level between clinical diagnoses by senior psychologist in the hospital and diagnoses/screening on Strengths and Difficulties Questionnaire conducted by parents. Using International Classification of Diseases (ICD) 10 (Indonesian version) as the gold standard, clinical child psychologists diagnosed 253 male and female elementary school children aged 7-14 years old. Parents of the children were requested to fill-out the SDQ questionnaire (SDQ-PR). Psychometric property of SDQ-PR was analyzed using test retest and Principal Axis Factoring Analysis. Screening quality of SDQ-PR was examined using Receiver Operating Characteristic (ROC), Likelihood Ratio (LR+ and LR-), and Chi square. Test retest reliability of SDQ-PR in all subscales were ($r = 0.562$ to $r = 0.731$) except subscale of peer-problem ($r = 0.174$). Chi square score indicated significant correlation between SDQ-PR and the diagnoses from child psychologist for hyperkinetic and behavior disorder, but not for emotional problem. This study concluded that out of 5 original subscales of SDQ-PR, it was revealed that only 3 can be used for Indonesian children, filled out by parents.

Keywords: child mental health; Indonesia; screening instrument; Strengths and Difficulties Questionnaire Parent Report (SDQ-PR).

Report from Sleman District Health Office of Yogyakarta Special Province, Indonesia indicates that children aged up to 19 years old were diagnosed suffer from behavioral and emotional disorders (46.37 %), developmental disorder (28.81%), neurosis, somatoform, and stress 16.50 %), intellectual disability disorder (3.78%), behavioral symptom related with physiological factor (3.68 %), mild depressive episode (0.52 %), and personality problems (0.31 %). Psychological epidemiology in childhood is becoming new research field and has become useful for child psychologist. As children's psychological problems grow in numbers, they are not

followed by the readiness of screening instruments to help child psychologists in their clinics. In Western countries the Rutter and the Achenbach questionnaires have been the most used research instruments for psychological examination in childhood. However, new instruments for symptom assessment in childhood are being used, among them the Strengths and Difficulties Questionnaires (SDQ). SDQ is a screening instrument for the investigation of the mental health of children and adolescents and was developed by Robert Goodman in 1997. Goodman's revision to Rutter's questionnaire by renewing and adding the content on the strength in children, had lead to the development of *Strengths and Difficulties Questionnaire*

¹ Address for correspondence: supra@ugm.ac.id

(SDQ). It is composed of 25 items subdivided in five subscales of five items each, measuring hyperactivity, emotional symptoms, conduct problem symptoms, interpersonal relationships and pro-social behavior. The SDQ can be filled out by parents (*SDQ Parent Report*), teachers (*SDQ Teacher Report*), and children (*SDQ Self Report*) (Goodman, 1994; Goodman, Meltzer, & Bailey, 1998). Studies on SDQ applications across countries are abundant in Western countries, but none or rare in Asia including Indonesia. Scoring of SDQ can be conducted through calculating the total score of each subscale or calculating the total difficulties score which is the sum of 4 subscales (emotional disorder, behavioral disorder, hyperactive-inattentive, and peer-problem).

Primary health centers need short, quick and easy to use instrument to assess psychological problems, even though this is not without critiques (Anthony & Barlow, 2002). Preliminary assessment using SDQ is useful to back-up the more accurate diagnosis in the Public Health Center level. SDQ provides information about brief descriptions focused on the strengths and difficulties experienced by child and adolescent (Black, Pulford, Christie, & Wheeler, 2010). In addition, SDQ also can be utilized by other professionals with or without background in mental health. These other professionals (nurse, midwife, general medical doctor) can act as the *gate keepers* to screen those children with psychological problems in Public Health Centers.

Psychometric properties of SDQ have been investigated. Results of Factor Analysis of SDQ for teachers, parents, and child self report indicated that this instrument was constructed by five factors to tap psychopathology and individual strengths. Internal consistency and test-

retest analysis was reported as satisfactorily. However, caution should be taken about number of sample and the interval between test-retest. Intercorrelation among SDQ scores from teachers, parents, and child self report was reported as moderate. SDQ is also able to differentiate children with or without psychopathology (Muris, Meesters, & van den Berg, 2003).

In the US and Europe, the most utilized instrument to measure psychopathological symptoms in child and adolescent is the Rutter and the Achenbach. Rutter Questionnaire is a short scale that are filled out by parents and teachers and shows its robust reliability and validity in various context of child psychopathology. However, this questionnaire is out of date and can not tap the condition such as concentration, impulsiveness, victimization, and pro-social behavior. On the other hand, Achenbach is considered suitable to tap the child mental health since the questionnaire is filled out by parents (Child Behavior Checklist/CBCL), teachers (Teacher Report Form/TRF) and the child (Youth Self Report/YSR). CBCL is accurate for clinical diagnoses of mental health among children and adolescents, but is less useful for screening or research because it is too long and contains items that irrelevant for majority of children (Muris, Meesters, & van den Berg, 2003).

SDQ has been translated into more than 60 languages including Bahasa Indonesia (Indonesian language), and free to be used for non commercial purposes. Initial research on SDQ in England by Goodman (1997) revealed that SDQ Parent Report (SDQ-PR) and SDQ Teacher Report (SDQ-TR) share similar function with Rutter questionnaire. SDQ Self Report (SDQ-SR) correlated well with both SDQ-PR and SDQ-TR (Goodman et.al, 1998).

Prediction of SDQ having 5 factors (emotional, conduct, hyperactivity-inattention, peer problem and pro-social) was also confirmed. Reliability coefficients calculated with internal consistency revealed *Cronbach Alpha* = 0.73, and correlation across informants = 0.34, and test retest reliability after 4-6 months is 0.62 (Goodman, 2001).

Studies on SDQ-TR, SDQ-PR and SDQ-SR have been conducted in other countries outside England. In Sweden, internal consistency of SDQ-PR was satisfactorily for all the subscales ($\alpha = 0.67 - 0.87$) except behavior disorders subscale ($\alpha = 0.52$). In England, SDQ-PR and SDQ-TR significantly correlated with independent clinical diagnoses (Goodman, Renfrew, & Mullick, 2000). Studies of SDQ in Germany showed that all subscales reliable significantly ($\alpha = 0.72 - 0.81$ for SDQ-PR, $\alpha = 0.75 - 0.83$ for SDQ-TR). All subscales of SDQ also correlated significantly with CBCL ($\alpha = 0.72 - 0.83$) and with TRF ($\alpha = 0.75 - 0.83$). Still in Germany, utilizing the total score or subscale scores of SDQ-PR, SDQ-TR and SDQ-SR were effective in predicting each clinical symptom in CBCL, TRF and YSR (Woerner, Becker, & Rothenberger, 2004; Becker, Woerner, Hasselhorn, Banaschewski, & Rothenberger, 2004).

Study of SDQ-PR and SDQ-SR in the Netherlands showed confirmation the existence of five factors in SDQ. Internal consistency for all subscales were satisfactorily, with SDQ-PR *Cronbach's Alpha* = 0.7 and 0.64 for SDQ-SR. The stability of test-retest showed the interclass correlation = 0.7 and above. Substantial correlation between total score of the disorders in SDQ and total score of CBCL ($r = 0.7$) (Muris et al., 2003). The internal consistency of SDQ-TR was considered as good with all α subscales as 0.8 (van

Widenfelt, Goedhart, Treffers, & Goodman, 2003) (Mieloo et al., 2012).

Studies of SDQ in Australia showed the reliability of SDQ-PR for all subscales as moderate to high ($\alpha = 0.59 - 0.8$) (Hawes & Daads, 2004). Japanese SDQ-PR was proven moderate for all the subscales ($\alpha = 0.52 - 0.77$) (Matsuishi et al., 2008). Study of SDQ-PR in France revealed moderate internal consistency ($\alpha = 0.54 - 0.74$) and the cut-off similar to those in England and USA for all the subscales, except pro-social scale in which the cut-off score is lower than those in the England and USA studies (Shojaei, Wazana, Pitrou, & Kovess, 2009). Chinese SDQ showed moderate to high internal consistency ranging from $\alpha = 0.45 - 0.81$ for SDQ-PR and $\alpha = 0.55 - 0.84$ for SDQ-TR. Cut-off score for subscales disorders was 17 for SDQ-PR and 15 for SDQ-TR (Lai et al., 2010). Internal consistency of Spain SDQ revealed $\alpha = 0.64 - 0.83$ for SDQ-TR and $\alpha = 0.58 - 0.77$ for SDQ-PR (Rodríguez-Hemández et al., 2012). It was reported that SDQ-PR in Denmark showed high internal consistency for all disorder subscales which is $\alpha \geq 0.7$ (Nielsen et al., 2012).

In Indonesia, SDQ had been translated and adapted by Tjhin Wiguna and Yohana Hestyanti (2012) and has been widely used for research and clinical practices. However, there is no clear investigation on the psychometric properties of it. Chen (2009) suggested that adapted instrument should be clearly known about the reliability and validity of the test. The scoring guideline of Indonesian SDQ that is provided in www.sdq.info.org is using English norm. As indicated by many studies of SDQ in various countries, the cut-off scores in different countries is influenced by its local culture. Report from Siregar and Wimbarti (2018) investigating SDQ-Teacher Report revealed diagnosis

consistency of hyperkinetic analyzed using receiver operating characteristic (ROC), likelihood ratio (LR) and Chi-Square indicated of ROC and LR cutoff point ≥ 7 , obtained prevalence of hyperkinetic disorder equals to 34.23% with sensitivity 72.4% and specificity 73.3%. Positive and Negative LR were 2.71 and 0.38. The Chi-Square analysis showed that there was a positive correlation between clinical psychologists' diagnosis and the SDQ-TR screening result. The current research investigating the SDQ-Parent Report.

This study sought to response the need of Indonesian Public Health Centers, especially in the remote areas where mental health professionals are rare. SDQ as a screening instrument of mental health was very much needed by local Health Centers in Indonesia. The study aimed at adapting, calculating the psychometric properties, the screening quality and the cut off scores of SDQ-PR for behavioral and emotional disorders, namely hyperkinetic disorder (F90), behavioral disorder (F91) and emotional disorder in childhood period (F93). Assessment on the diagnostic quality of an instrument needs a gold standard for diagnosis (Dahlan, 2001), and Aboraya, France, Young, Curci, & LePage (2005) defines gold standard diagnosis as the standard for diagnosis that make use all available validity criteria. Clinical psychologist should utilize all available validity criteria to achieve most accurate diagnoses. The gold standard used in this research was the diagnosis from senior psychologist from public health center based on ICD 10 Indonesian version for behavioral and emotional problem that comprise hyperkinetic disorder (F90), behavioral problem (F91) and emotional problem which occurs in childhood (F93).

Method

Subjects of this research were 253 elementary school boys and girls, age 7-14 years old in Sleman District, Yogyakarta Special Province of Indonesia. All of them had been in that school for at least a year. Informed consent was given by parents of children as respondent of this research.

Procedure

The first step that we did was translating the SDQ PR from English to Bahasa Indonesia (Indonesian language) and backtranslated again into English as suggested in Hambleton, Merenda, & Spielberger (2005) as the rule of adapting an instrument to other culture. Translation from English to Bahasa Indonesia was done by two English native speakers who can speak Bahasa Indonesia have been living in Yogyakarta several years. They are English language teachers Original SDQ-PR (English version) was translated to SDQ-PR Indonesian version and then was evaluated by three senior psychologists who are competent in cross-cultural child psychology. The Indonesian version was then backtranslated into English to see the consistency of meaning. The backtranslation was done by two Indonesian English teachers. Professional psychologist involved in this study to assure that the translation meets the requirement of psychological construct of the instrument. Cognitive debriefing was given to 20 parents of elementary school children in Sleman District of Yogyakarta. This step was taken to assure that they comprehended SDQ-PR Indonesia being studied (Plougman, Austin, Stefanelli, Godwin, 2010). The next step was videotaping sample behaviors of children who show symptoms of hyperkinetic disorder, behavioral disorder, and emotional disorder occur during childhood, all according to

ICD 10 Indonesian version. This video was utilized to test the screening quality of the psychologists who would collect the data. All symptoms showed in the video were validated by a senior child psychologist. The next step was conducting training for junior psychologists who would collect data. The training was given by a senior child psychologist by utilizing the ICD 10 Indonesian version. The training module was created by the current researchers. The aim of the training was to create the same perception on diagnosing hyperkinetic disorder, behavioral disorder, and emotional disorder occur during childhood. The last step was testing the diagnostic quality of the psychologists who would be the data collector. Each psychologist data collector rated the 3 videos containing the disorders. The quality indicator was Kappa score from interrater reliability analysis by comparing ratings by data collector psychologist and the senior child psychologist as the gold standard. For all the data collector psychologists, Kappa scores range from 0.71 to 1.

Data analysis

Data analysis were three steps. First, psychometric property analysis which were examining the reliability through Test Retest and the construct validity through Principal Axis Factoring. Second, applying Receiver Operating Curve (ROC) and Likelihood Ratio (LR) to examine the quality of screening. Third, to examine the

correlation between SDQ-PR scores with diagnostic score done by the data collector psychologists by calculating the Chi square and Contingency Coefficient.

Result

Participants of this study were 253 elementary school children in which 24.5% (62) were females and 75.5% (191) were males. As for the ages of the participants, is depicted on Table 1. As for the school grade, two participants were at grade 1; 54 were at grade 2; 58 at grade 3; 48 were at grade 4; 46 were at grade 5; and 45 were at grade 6.

Table 2 depicts minimal, maximal, and the mean scores for Hyperactivity/Inattention, Emotional Disorder, Behavioral Disorder, Problem with Peer, and Pro-social Behavior.

Table 1
Age of Participants

Age	Total
7	21
8	50
9	49
10	42
11	36
12	38
13	13
14	4

Tabel 2
Mean Score of SDQ-PR Subscales

	Hyperactivity/ Inattention	Emotional Disorder	Behavioral Disorder	Problem with Peer	Pro-social Behavior
Xmax	10	10	10	8	10
Xmin	0	0	0	0	1
Mean	4.439	2.862	2.616	2.685	5.517

Table 3 shows the Demographic Data and Diagnosis by Psychologist. Hyperkinetic disorder mostly happened to grade 2 participants, behavioral disorder at grade 3, and emotional disorder at grade 2 and grade 6.

Reliability of SDQ-PR

Initially, the reliability was analyzed using Cronbach alpha, but the number of item was too short, and facing threat to the reliability. Test retest then applied and showed of hyperkinetic $r = 0.731$; for emotional disorder $r = 0.684$; for behavioral disorder $r = 0.562$; for peer problem $r = 0.174$; and for prosocial behavior $r = 0.647$. This indicating that reliability of Peer Problem was not sufficient, while the rest were moderate.

Construct validity of SDQ-PR

Construct validity examination using Principal Axis Factoring to SDQ-PR Indonesia resulting in 8 factors (See Table 4). Factor 1 consisted of item number 2, 5, 10, 12, 15, 18, 22, and 23, fit with hyperkinetic factor. Factor 2 consisted of item number 3, 6, 8, 13, 16, and 24, fit with emotional problems factor. Factor 3 consisted of item number 4, 7, 9, 11, 14, and 20, fits with pro-social factor. Factor 4 consisted of item number 1 and 25, fit with empathy ability. Factor 5 consisted of item number 21 which fits with inattention.

Factor 6 consisted of item number 17 which fits with kindness to kids. Factor 7 consisted of item number 2 and 16 which fit with "worries". Factor 8 consisted of item number 19 which fit with "peer problem".

SDQ-PR has 5 factors namely emotional problem, behavior problem, hyperactivity-inattention (hyperkinetic), peer problem and pro-social. Table 4 indicates that only 3 factors confirmed in this current study, those were hyperkinetic, emotional problems, and pro-social. The other 2 factors were spread-out into 5 different factors. In this research behavior and peer-problem did not form as solid factors but are broken into 5 smaller factors. This might be due to parents read the items of behavior and peer-problems as either emotional problem or hyperkinetic factors. Three factors that were valid comprise of 20 items, the remaining 5 items could not be used to assess problems as intended in the original SDQ tool. In short, for Indonesian usage, only 3 factors are valid and can be utilized for clinical practice.

The total variance explained by the eight factors was 58.307%. Item number 2 was grouped into two factors which were factor 1 and factor 7, item number 16 was grouped into factor 2 and 7. Table 4 shows the result of principal axis factoring rotation.

Screening Quality of SDQ-PR

ROC analysis revealed the Area Under Curve (AUC) score of 70.6 % (95% IK, 63.4%-77.7%), $p < 0.001$ for screening of hyperkinetic disorder, 70.0 % (95% IK, 59.7% - 80.3%), $p < 0.001$ for screening of behavioral disorder, and 58.8% (95% IK, 42.9% - 74.7%), $p < 0.001$ for screening of emotional problem.

Table 3
Demographic Data and Diagnosis by Psychologist Pindah di halaman 6

Criteria		Hiperkinetic Disorder	Behavioral Disorder	Emotional Disorder
Age	7	6 (2.6%)	1 (0.4%)	3 (1.3%)
	8	21 (9.1%)	5 (2.2%)	2 (0.9%)
	9	14 (6.0%)	10 (4.3%)	3 (1.3%)
	10	12 (5.2%)	6 (2.6%)	3 (1.3%)
	11	6 (2.6%)	3 (1.3%)	3 (1.3%)
	12	10 (4.3%)	4 (1.7%)	2 (0.9%)
	13	4 (1.7%)	1 (0.4%)	2 (0.9%)
	14	1 (0.4%)	1 (0.4%)	0 (0%)
School Grade	1	1 (0.4%)	0 (0%)	0 (0%)
	2	25 (10.8%)	6 (2.6%)	5 (2.2%)
	3	17 (7.3%)	12 (5.2%)	3 (1.3%)
	4	12 (5.2%)	5 (2.2%)	3 (1.3%)
	5	7 (3.0%)	3 (1.3%)	2 (0.9%)
	6	12 (5.2%)	5 (2.2%)	5 (2.2%)
Sex	Male	70 (30.2%)	27 (11.6%)	11 (4.7%)
	Female	4 (1.7%)	4 (1.7%)	7 (3.0%)

Table 4
Result of Principal Axis Factoring Rotation

Item	Factor							
	1	2	3	4	5	6	7	8
SDQP12	.696	.191	-.153	.086	-.042	-.059	-.115	-.048
SDQP10	.615	-.019	-.117	-.106	.101	-.054	.159	.067
SDQP18	.569	.102	-.084	-.082	.031	-.111	-.121	.055
SDQP2	.514	-.049	.058	-.071	.022	-.065	.512	.014
SDQP15	.484	.232	-.033	-.130	.276	.221	.052	-.121
SDQP5	.440	.273	-.053	.046	-.022	-.038	.185	.179
SDQP22	.342	.107	-.062	.112	.180	-.338	.050	.059
SDQP23	.281	.061	.162	.061	.026	.159	.110	.025
SDQP13	.187	.609	.029	-.039	-.053	.020	-.199	-.029
SDQP3	.101	.576	.076	.018	-.023	-.229	.044	-.096
SDQP24	.161	.542	.050	-.004	.073	.092	-.022	.055
SDQP16	.091	.542	-.090	.053	.076	.315	.321	-.090
SDQP8	-.004	.512	.069	.033	-.090	-.002	.010	.247
SDQP6	.025	.415	-.062	.092	-.129	-.196	.066	.119
SDQP14	.081	.125	-.684	-.112	-.097	-.044	-.003	.231
SDQP20	-.097	.026	.612	.156	-.064	.020	-.121	.191
SDQP4	-.121	.064	.521	.068	.065	-.038	.146	.150
SDQP7	.002	.027	-.434	.059	.133	-.135	.058	-.062
SDQP11	-.007	-.133	-.431	.031	.093	-.021	-.094	.133
SDQP9	.023	.067	.421	.304	-.142	.090	-.106	.079
SDQP1	-.035	.070	.162	.857	-.087	-.010	.032	-.058
SDQP25	.060	.012	-.092	-.244	.233	-.049	.058	-.182
SDQP21	.148	-.166	-.148	-.133	.750	-.034	-.005	.117
SDQP17	-.101	-.091	.418	.140	-.022	.570	-.018	.084
SDQP19	.261	.231	.084	-.013	.114	.015	.019	.434

Optimum cutting point of SDQ-PR based on ROC analysis to screen hyperkinetic disorder was 6, with sensitivity of 67.6% and specificity of 63.3%. It means a client with hyperactivity-inattention score ≥ 6 was a suspect of having hyperkinetic disorder, thereby he/she needed further clinical examination. Optimum cutting point of SDQ-PR based on ROC analysis to screen behavioral disorder was 4 with sensitivity of 70.96% and specificity of 69.15%. It meant a client with behavioral

disorder score ≥ 4 was a suspect of having behavioral disorder, thereby he/she needed further clinical examination. The optimum cutting point of SDQ-PR based on ROC analysis to screen emotional disorder was 4 with sensitivity of 50% and specificity of 53.27%. Sensitivity score of 50% was considered as low, therefore SDQ-PR for emotional disorder could not be used to screen emotional disorder occurred during. Result of ROC analysis is depicted on Table 5.

Tabel 5
Result of ROC Analysis

SDQ-PR	AUC score	Cut-off Score	Sensitivity	Specificity
Hyperkinetic	70.6 %	≥ 6	67.6%	63.3%
Behavioral disorder	70.0 %	≥ 4	70.96%	69.15%
Emotional disorder	58.8%	≥ 4	50%	53.27%

Likelihood ratio analysis

Based on Likelihood Ratio (LR) analysis, cutting score ≥ 6 on SDQ-PR subscale for hyperactivity-inattention, it was revealed that prevalence of hyperkinetic disorder was 31.89% with sensitivity of 67.6% and specificity of 63.3%. It meant, cutting score of ≥ 6 for SDQ-PR subscale hyperactivity-inattention was, (1) able to screen with positive results on 67.6% children who were suffering from hyperkinetic disorder, (2) able to screen with negative results on 63.3% children who were not suffering from hyperkinetic disorder.

Analysis of LR showed that LR+ value of 1.84 and LR- value of 0.51 apparent to all samples. It meant that children with hyperkinetic disorder had the probability of 1.84 times higher to have a positive screening results on SDQ-PR subscale hyperactivity-inattention than those children who do not have hyperkinetic disorder. In addition, children who were with hyperkinetic disorder had a probability of 0.51 times lower to be screened with SDQ-PR subscale hyperactivity-inattention with negative result compared to those who were not with hyperkinetic disorder.

Fagan Likelihood-Ratio Nomogram showed if a child had a positive result based on SDQ-PR subscale hyperactivity-inattention, the tendency that he/she would also suffer from hyperkinetic disorder increases from 32% to 48%. If the child had negative result on SDQ-PR subscale hyperactivity-inattention, the tendency that he/she would suffer from

hyperkinetic disorder decreased from 32 to 20%.

Based on *Likelihood Ratio* (LR) analysis, with cutting score ≥ 4 on SDQ-PR subscale of behavior disorder, it was revealed that prevalence of behavior disorder is 13.36% with sensitivity of 71% and specificity of 69.2%. It meant that cutting score of ≥ 4 on SDQ-PR subscale of behavior disorder was, (1) able to screen with positive results on 71% children who were suffering from behavioral disorder, (2) able to screen with negative results on 69.2% children who were not suffering from behavior disorder.

Analysis of LR showed that LR+ value of 2.3 and LR- value of 0.42 apparent to all samples. It meant that children with behavior disorder had the probability of 2.3 times higher to have a positive screening results on SDQ-PR subscale behavior disorder than those children who do not have behavior disorder. In addition, children who were with behavioral disorder had a probability of 0.42 times lower to be screened with SDQ-PR subscale behavior disorder with negative result compared to those who were not with behavior disorder.

Fagan Likelihood-Ratio Nomogram showed if a child had a positive result based on SDQ-PR subscale of behavior disorder, the tendency that he/she would also suffer from behavior disorder increased from 13% to 25%. If the child has negative result on SDQ-PR subscale of behavior disorder the tendency that he/she would suffer from behavior disorder decreased from 13% to 7%.

Based on Likelihood Ratio (LR) analysis, cutting score ≥ 4 on SDQ-PR subscale for emotional disorder, it was revealed that prevalence of emotional disorder was 7.75% with sensitivity of 50% and specificity of 53.3%. It meant that cutting score of ≥ 4 for SDQ-PR subscale emotional disorder was, (1) able to screen with positive results on 50% children who were suffering from emotional disorder, (2) able to screen with negative results on 53.3% children who were not suffering from emotional disorder.

Analysis of LR showed that LR+ value of 1.07 and LR- value of 0.94 apparent to all samples. It meant that children with emotional disorder have the probability of 1.07 times higher to have a positive screening results on SDQ-PR subscale of emotional disorder than those children who did not have emotional disorder. In addition, children who were with emotional disorder had a probability of 0.94 times lower to be screened with SDQ-PR subscale of emotional disorder with negative result compared to those who were not with emotional disorder.

Fagan Likelihood-Ratio Nomogram showed if a child had a positive result based on SDQ-PR subscale of emotional disorder, the tendency that he/she would also suffer from emotional disorder increased from 7.75% to 8%. If the child had negative result on SDQ-PR subscale of emotional disorder, the tendency that he/she would suffer from hyperkinetic disorder decreased from 7.75 %to 7.5%.

Complete LR analysis is shown on Table 6.

Chi square analysis

Result of *Chi-Square* analysis showed a significant correlation between SDQ-PR screening and diagnoses made by psychologist for hyperkinetic disorder

(Pearson $\lambda^2 = 19.288$, $p < 0.05$), coefficient contingency = 0.277. Coefficient contingency showed that this correlation was not so strong.

Table 6
Result of LR Analysis

Disorder	Prevalence	LR+	LR-
Hyperkinetic	31.89%	1.84	0.51
Behavioral	13.36%	2.3	0.42
Emotional	7.75%	1.07	0.94

Result of *Chi-Square* analysis showed a significant correlation between SDQ-PR screening and diagnoses made by psychologist for behavioral disorder (Pearson $\lambda^2 = 18.718$, $p < 0.05$), coefficient contingency = 0.273. This *coefficient contingency* was also not significant to indicate the correlation between SDQ-PR with the diagnoses made by psychologist.

Result of *Chi-Square* analysis showed no significant correlation between SDQ-PR screening and diagnoses made by psychologist for emotional disorder (Pearson $\lambda^2 = 0.071$, $p > 0.05$) coefficient contingency = 0.018. This *coefficient contingency* was also not significant to indicate the correlation between SDQ-PR with the diagnoses made by psychologist. Complete *Chi Square* analysis is shown in Table 7.

Discussion

Test retest reliability on all the subscale of SDQ-PR showed inconsistent findings. Reliability coefficient for hyperactivity-inattention was $r = 0.731$, for emotional disorder $r = 0.684$, for behavioral disorder was $r = 0.562$, for peer-problem was $r = 0.174$, and for social competence was $r = 0.647$. Except for peer-problem, the other four aspects showed reliability of moderate in nature. This finding was

Table 7
Result of Chi Square Analysis

Disorder	<i>Pearson Chi square</i>	Sig. level	<i>Coefficient contingency</i>
Hyperkinetic	19.288	< 0.05	0.277
Behavioral	18.718	< 0.05	0.273
Emotional	0.071	0.05	0.018

different from the previous studies from different cultural background. In Sweden, for example, SDQ-PR showed better psychometric property, in which the Cronbach Alpha ranged from 0.67 – 0.87, except for behavioral problem ($\alpha = 0.52$) (Smedje, Broman, Hetta, & Knorrning, 1999; Malmberg, Rydell, & Smedje, 2003). SDQ-PR Germany showed satisfactorily reliability for all subscales, α ranged from 0.72 to 0.8 (Woerner, Becker, & Rothenberger, 2004; Becker, Hagenberg, Roessner, Woerner, & Rothenberger, 2004). The fact was also true for SDQ-PR Australia which showed satisfactorily reliability for all subscales with α ranges from 0.59 to 0.8 (Hawes & Daads, 2003). SDQ-PR Japan also had satisfactorily reliability for all the subscales ($\alpha = 0.52$ to 0.77) (Matsuihi et.al., 2008). The true was for SDQ-PR France ($\alpha = 0.54$ to 0.74) (Shojaei, Wazana, Pitrou, & Kovess, 2009), SDQ-PR China ($\alpha = 0.45$ to 0.81) (Lai et al., 2010) and SDQ-PR Spain ($\alpha = 0.58$ to 0.77) (Rodríguez-Hemández et al., 2012). Knowing the reliability of SDQ-PR from various countries, it is apparent that Indonesian SDQ-PR is less satisfactorily.

This study found that SDQ-PR Indonesian version consisted only 3 factors instead of 5 factors as constructed in the original SDQ-PR. The 3 factors were Hyperkinetic, Emotional Problems, and Pro-social. The other 2 which were Behavior problems and Peer-problem did not show good reliability. The eight factors described the variance of 58.307% from the total variance. There were two items

grouped into two different factors. Item number 2 was grouped under factor 1 (Hyperactivity) and factor 7 (Worries). Item number 16 was grouped under factor 2 (Emotional disorder) and factor 7 (Worries). This happened due to heavy factor loading of item number 2 and 16 in the two factors.

As far as we know, there is no research on SDQ-PR psychometric properties that formed up to 8 factors. Most of the factor analysis in the other studies confirmed the 5 factors of SDQ-PR. Goodman research (2001) SDQ-PR confirmed 5 factors (Emotional, Conduct, Hyperactivity-inattention, Peer problem and Prosocial). In Germany, The Netherland, and Japan, SDQ-PR also confirmed 5 factors, similar to the 5 factors from original SDQ-PR version in England (Muris, Meesters, & van den Berg, 2003; Woerner, Becker, & Rohenberger, 2004; Matsuishi, et.al, 2008).

Only moderate value of reliability and the forming of SDQ-PR Indonesia into 8 factors might be due to the adaptation procedures of this tool. Some aspects should be considered at this point are (1) building the equity of concept, function, and operationalization of the concept/theory utilized for building the SDQ-PR. Building the equity can be done through professional judgement to ensure that SDQ-PR is conceptually appropriate to be adapted into Bahasa Indonesia (Indonesian language), Construction of SDQ-PR should be ensured to have the same goal, and can be operationalized

using the same way as the original. In this current research, these procedures involving 3 senior psychologists for professional judgment. Two of them were child psychologists and one of them was cross cultural psychologist who was fluent in English language and comprehended Indonesian and English cultures.

In addition to professional judgment, test adaptation should also make sure that the test format, and administration is equal in the two cultures (England and Indonesia). The fact that educational background of the parents in this research varied, and generally low to moderate education, and there were even several parents can not write nor read, caution has to be exercised. Report also gathered some parents said that they were not used to paper and pencil works. Further researchers should consider these minor issues in Indonesia.

Based on ROC analysis it was revealed that SDQ-PR quality for screening of emotional disorder is moderately low with sensitivity of 50 %. Sensitivity value of 50 % is a probability, that is why it can also be concluded that SDQ-PR Indonesia Emotional disorder subscale is not sensitive enough to screen Emotional disorder in children by the parents. This is aligned with previous research in Australia that revealed low value of sensitivity for Emotional disorder subscale to screen emotional problems, which is 36.0% (Mathai, Anderson, & Brown, 2004). In Sweden, however, SDQ-PR Sweden had a sensitivity value of 72.95% for emotional disorder subscale (Malmberg, Rydell, & Smedje, 2003).

Hyperactivity-inattention and behavioral disorder subscales showed sensitivity values higher than that of emotional disorder subscale. This finding aligns with the previous findings from

other cultural background, such as the USA, Australia, England, Sweden and Bangladesh. Research of SDQ-PR in England showed sensitivity of emotional disorder subscale is highest on SDQ-SR. The findings then, suggested that parents and teachers were not recommended to screen for emotional disorder but they were good for hyperactivity and behavioral disorder.

LR+ and LR- as well as Fagan Likelihood Nomogram analysis were never been applied to any SDQ studies in the past, but information scores depicted from LR is beneficial for mental health practitioners in the field. Behavioral disorder subscale was the best in LR+ and LR- (LR+ = 2.3, LR- = 0.42) compared to Hyperactivity-inattention subscale (LR+ = 1.84, LR- = 0.51) and emotional disorder subscale (LR+ = 1.07, LR- = 0.94). However Fagan Likelihood Nomogram showed increment the highest value of probability to be diagnosed if an individual was screened positive for Hyperactivity was 16% and decrement the highest value of probability to be diagnosed if an individual was screened negative for Hyperactivity was 12%. This result showed that Hyperactivity-Inattention subscale and Behavioral disorder subscale were strong screening factors that could predict mental health problems in children especially those who came to seek help in the public health centers in Indonesia. Result of Chi square analysis also confirmed that SDQ-PR significantly predicted behavioral and hyperkinetic disorders, but not emotional disorder. To conclude, SDQ-PR Indonesia is sufficient for screening behavioral disorder and hyperactivity-inattention disorder.

Conclusion

This study aimed at examining the psychometric property and agreement level between clinical diagnoses by senior psychologist in the hospital and diagnoses/screening on Strengths and Difficulties Questionnaire conducted by parents. Out of 5 original subscales of SDQ-PR, it was revealed that only 3 can be used for Indonesian children, filled out by parents.

Recommendation

Compared to the research findings on SDQ-PR in other country, our research had comparable number of subjects (ranges from 200-300 children), but we found different results especially in the number of factors consisted the SDQ-PR. We suggest that future research should consider to revisit the theoretical background of SDQ-PR to understand the relevance of the criteria used to determine each of the problems in SDQ-PR. It is also suggested that the assessor psychologist to be re-assessed on their competence and consistency in determining the diagnosis.

Acknowledgement

Special thanks go to Lembaga Penelitian dan Pengabdian Masyarakat UGM for awarding the grant to the researchers. Thanks also go to schools, parents, and senior psychologist Dr. Gamayanti who involved in this research.

References

- Aboraya, A., France, C., Young, J., Curci, K., & LePage, J. (2005). The validity of psychiatric diagnosis revisited: The clinician's guide to improve the validity of psychiatric diagnosis. *Psychiatric Epidemiology*, 2(9), 48-55.
- Antony, M. M., & Barlow, D. H. (2002). *Handbook of assessment and treatment planning for psychological disorders*. Los Angeles: The Guilford Press.
- Becker, A., Woerner, W., Hasselhorn, M., Banaschewski, T., & Rothenberger, A. (2004). Validation of the parent and teacher SDQ in a clinical sample. *European Child and Adolescent Psychiatry*, 13, 11-16. doi: [10.1007/s00787-004-2003-5](https://doi.org/10.1007/s00787-004-2003-5)
- Becker, A., Hagenberg, N., Roessner, V., Woerner, W., & Rothenberger, A. (2004). Evaluation of the self-reported SDQ in a clinical setting: Do self-reports tell us more than ratings by adult informants? *European Child and Adolescent Psychiatry*, 13, 17-24.
- Black, S., Pulford, J., Christie, G., & Wheeler, A. (2010). Differences in New Zealand secondary school students' reported strengths and difficulties. *New Zealand Journal of Psychology*, 39(1), 19-23.
- Chen, W. S. (2009). Questionnaire translation and psychometric properties evaluation. *SEGi Review*, 2(2), 62-71.
- Dahlan, M. S. (2001). *Penelitian diagnostik: Dasar-dasar teoritis dan aplikasi dengan program SPSS dan stata*. Jakarta: Penerbit Salemba Medika.
- Goodman, R. (1994). A modified version of the rutter parent questionnaire including extra items on children's strengths: A research note. *Journal of Child Psychology and Psychiatry*, 35, 1483-1494.
- Goodman, R. (1997). The strength and difficulties questionnaires: A research not. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaires: A pilot study on the

- validity of the self report version. *European Child and Adolescent Psychiatry*, 7, 125-130.
- Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strength and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child and Adolescent Psychiatry*, 9, 124-134.
- Goodman, R. (2001). Psychometric properties of the Strength and Difficulties Questionnaire (SDQ) [Abstrak]. *Journal of American Academy of Child and Adolescent Psychiatry*, 40, 1337-1345.
- Hambleton, R. H. Merenda, P. F. & Spielberger, C. D. (2005). *Adapting educational and psychological test for cross-cultural assessment*. New Jersey: Lawrence Erlbaum Associates.
- Hawes, D. J., & Dadds, M. R. (2004). Australian data and psychometric properties of the Strength and Difficulties Questionnaire (SDQ). *Australian and New Zealand Journal of Psychiatry*, 38, 644-65.
- Lai, K. Y. C., Luk, E. S. L., Leung, P. W. L., Wong, A. S. Y., Law, L., & Ho, K. (2010). Validation of the Chinese Version of the Strengths and Difficulties Questionnaire in Hong Kong. *Social Psychiatry Epidemiology*, 45, 1179-1186. doi: [10.1007/s00127-009-0152-z](https://doi.org/10.1007/s00127-009-0152-z)
- Malmberg, M., Rydell, A. M., & Smedje, H. (2003). Validity of Swedish version of the Strength and Difficulties Questionnaire (SDQ-Swe). *Nordic Journal Psychiatry*, 57, 357-363. doi: [10.1080/08039480310002697](https://doi.org/10.1080/08039480310002697)
- Mathai, J., Anderson, P., & Bourne, A. (2004). Comparing psychiatric diagnoses generated by the Strengths and Difficulties Questionnaire with diagnoses made by clinicians. *Australian and New Zealand Journal of Psychiatry*, 38, 639-643
- Matsuishi, T., Nagano, M., Araki, Y., Tanaka, Y., Iwasaki, M., Yamashita, Y., Nagamitsu, S., Iizuka, C., Ohya, T., Shibuya, K., Hara, M., Matsuda, K., Tsuda, A., & Kakuma, T. (2008). Scale properties of the Japanese version of the Strengths and Difficulties Questionnaire (SDQ): A study of infant and school children in community samples. *Brain and Development*. 30, 410-415. doi: [10.1016/j.braindev.2007.12.003](https://doi.org/10.1016/j.braindev.2007.12.003)
- Mieloo, C., Raat, H., Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and Reliability of the Strengths and Difficulties Questionnaire in 5-6 Year Olds: Differences by Gender or by Parental Education? *Plos One*. 7, 1-8.
- Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry*, 12, 1-8. doi: [10.1007/s00787-003-0298-2](https://doi.org/10.1007/s00787-003-0298-2)
- Niclasen, J., Teasdale, T. W., Andersen, A. N., Skovgaard, A. M., Elberling, H., & Obel, C. (2012). Psychometric Properties of the Danish Strength and Difficulties Questionnaire: The SDQ assessed for more than 70,000 raters in four different cohorts. *Plos One*. 7, 1-8.
- Rodríguez-Hemández, P. J., Betancort, M., Ramírez-Santana, G. M., García, R., Sanz-Álvarez, E. J., & Cuevas-Castresana, C. D. (2012). Psychometric properties of the parent and teacher versions of the Strength and Difficulties Questionnaire (SDQ) in a Spanish

- sample. *International Journal of Clinical and Health Psychology*, 12, 265-279.
- Shojaei, T., Wazana, A., Pitrou, I., & Kovess, V. (2009). The Strengths and Difficulties Questionnaire: Validation study in french school-aged children and cross-cultural comparisons. *Social Psychiatry and Psychiatric Epidemiology*, 44, 740-747. doi: [10.1007/s00127-008-0489-8](https://doi.org/10.1007/s00127-008-0489-8)
- Siregar, J & Wimbarti, S. (2018). Screening for the hyperkinetic disorder by using Strength and Difficulties questionnaire Teacher-Report (SDQ-TR) in Indonesian school-aged children. *Global Journal of Psychology Research: New Trends and Issues*, 8(2), 76-87.
- Smedje, H., Broman, J. E., Hetta, J., & Knorrning, A. L. (1999). Psychometric properties of swedish version of the Strengths and Difficulties Questionnaire. *European Child and Adolescent Psychiatry*, 8, 63-70
- van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry*, 12, 281-289. doi: [10.1007/s00787-003-0341-3](https://doi.org/10.1007/s00787-003-0341-3)
- Woerner, W., Becker, A., & Rohenberger, A. (2004). Normative data and scale properties of the German parent SDQ. *European Child and Adolescent Psychiatry*, 13, 3-10. doi: [10.1007/s00787-004-2002-6](https://doi.org/10.1007/s00787-004-2002-6)
- Wiguna, T. & Hestyanti, Y. (2012). *What is the SDQ?*. Retrieved from <https://www.sdqinfo.com/a0.html>